Sequence Listing

SEQ ID NO.: 1: ESX cDNA ORF and deduced amino acid sequence
(See Figure 1)

SEQ ID NO.: 2: ESX cDNA sequence (5' untranslated + ORF + 3'
untranslated)1907 b.p.

```
cggccagatacctcagcgctacctggcggaactggatttctctcccgcctgccggcctgcct
gccacagccggactccgccactccggtagcctcatggctgcaacctgtgagattagcaacat
ttttagcaactacttcagtgcgatgtacagctcggaggactccaccctggcctctgttcccc
ctgctgccacctttggggccgatgacttggtactgaccctgagcaaccccagatgtcattg
gagggtacagagaaggccagctggttggggggaacagccccagttctggtcgaagacgcaggt
tctggactggatcagctaccaagtggagaagaacaagtacgacgcaagcgccattgacttct
cacgatgtgacatggatggcgccaccctctgcaattgtgcccttgaggagctgcgtctggtc
tttgggcctctggggggaccaactccatgcccagctgcgagacctcacttccagctcttctga
tgagctcagttggatcattgagctgctggagaaggatggcatggccttccaggaggccctag
acccagggcccttgaccagggcagccccttgcccaggagctgctggacgacggtcagcaa
gccagcccctaccaccccggcagctgtggcgcaggagcccccctccctggcagctctgacgt
ctccaccgcagggactggtgcttctcggagctcccactcctcagactccggtggaagtgacg
tggacctggatcccactgatggcaagctcttccccagcgatggttttcgtgactgcaagaag
ggggatcccaagcacgggaagcggaaacgaggccggccccgaaagctgagcaaagagtactg
ggactgtctcgagggcaagaagagcaagcacgcgcccagaggcacccacctgtggggagttca
tccgggacatcctcatccacccggagctcaacgagggcctcatgaagtgggagaatcggcat
gaaggcgtcttcaagttcctgcgctccgaggctgtggcccaactatggggccaaaagaaaaa
aacagcaacatgacctacgagaagctgagccgggccatgaggtactactacaaacgggaga
tcctggaacgggtggatggccggcgactcgtctacaagtttggcaaaaactcaagcggctgg
aaggaggaagaggttctccagagtcggaactgagggttggaactatacccgggaccaaactc
acggaccactcgaggcctgcaaaccttcctgggaggacaggcaggccagatggcccctccac
tggggaatgctcccagctgtgctgtggagagaagctgatgtttggtgtattgtcagccatc
gtccttggactcggagactatggcctcgcctcccccaccctcctcttggaattacaagccctg
gggtttgaagctgactttatagctgcaagtgtatctccttttatctggtgcctcctcaaacc
cagtctcagacacttaaatgcagacaacaccttcttcctgcagacacttggactgagccaag
gaggcttggggaggccctagggagcaccgtgatggagaggacagagcagggggctccagcactt
ctttctggactggcgttcacctccctgctcagtgcttgggctccacgggcagggggtcagagc
actccctaatttatgtgctatataaatatgtcagatgtacatagagatctattttttctaaa
acattcccctccccactcctctcccacagagtgctggactgttccaggccctccagtgggct
gatgctgggacccttaggatggggctcccagctcctttctcctgtgaatggaggcagagacc
tccaataaagtgccttctgggctttttctaaaaaaaaaaaaaaaaaa
```

SEQ ID NO.: 3: Complete ESX deduced amino acid sequence (see
Figure 1)

SEQ ID NO.: 4: First variable region (nucleotides 1-189 of
Figure 1)

```
atggctgcaacctgtgagattagcaacattttttagcaactacttcagtgcgatgtacagctc
ggaggactccaccctggcctctgttccccctgctgccacctttggggccgatgacttggtac
tgaccctgagcaaccccagatgtcattggagggtacagagaaggccagctggttggggggaa
cag
```
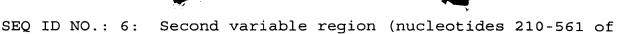
SEQ ID NO.: 5: ※ Pointed※ region (nucleotides 190-309 of
Figure 1)

```
ccccagttctggtcgaagacgcaggttctggactggatcagctaccaagtggagaagaacaa
gtacgacgcaagcgccattgacttctcacgatgtgacatggatggcgccaccctctgc
```

SEQ ID NO.: 6: Second variable region (nucleotides 210-561 of Figure 1)

aattgtgcccttgaggagctgcgtctggtctttgggcctctgggggaccaactccatgccca
gctgcgagacctcacttccagctcttctgatgagctcagttggatcattgagctgctggaga
aggatggcatggccttccaggaggccctagacccagggcccttttgaccagggcagcccttt
gcccaggagctgctggacgacggtcagcaagccagcccctaccaccccggcagctgtggcgc
agga

SEQ ID NO.: 7: Deduced amino acid sequence for second variable region (amino acids 104-187 of Figure 1)

asn-cys-ala . . . . gly-ala-gly

SEQ ID NO.: 8: Serine-rich region (nucleotides 562-714 of Figure 1)

gccccctcccctggcagctctgacgtctccaccgcagggactggtgcttctcggagctccca
ctcctcagactccggtggaagtgacgtggacctggatcccactgatggcaagctcttcccca
gcgatggttttcgtgactgcaagaagggg

SEQ ID NO.: 9: Third variable region (nucleotides 715-819 of Figure 1)

gatcccaagcacgggaagcggaaacgaggccggccccgaaagctgagcaaagagtactggga
ctgtctcgagggcaagaagagcaagcacgcgcccagaggcacc

SEQ ID NO.: 10: Ets DNA Binding domain (nucleotides 820-1062 of Figure 1)

cacctgtgggagttcatccgggacatcctcatccacccggagctcaacgagggcctcatgaa
gtgggagaatcggcatgaaggcgtcttcaagttcctgcgctccgaggctgtggcccaactat
ggggccaaaagaaaaagaacagcaacatgacctacgagaagctgagccgggccatgaggtac
tactacaaacgggagatcctggaacgggtggatggccggcgactcgtctacaagttt

SEQ ID NO.: 11: Fourth variable region (nucleotides 1063-1113 of Figure 1)

ggcaaaaactcaagcggctggaaggaggaagaggttctccagagtcggaac

SEQ ID NO.: 12: C-terminal 16 amino acids (amino acids 356-371 of Figure 1)

lys-asn-ser . . . ser-arg-asn

SEQ ID NO: 13 5'ESX-DBD

5'-CCGGGACATCCTCA TCCACCC-3'

SEQ ID NO: 14 3' ESX-DBD

5'-GTACCTCATGGCCCGGCTCAG-3' (SEQ ID NO. 14)).

SEQ ID NO: 15 Mouse ESX genomic sequence.

```
   1   GGATCCTTCC AAGGCACTGA CCTCACCCAA TTCTTTCTCA CTTTTCTCCT
  51   CCATTTAACT GTGGACGGAA TCAATACTCA GGGGGATGCG CTAGCTCTAA
 101   GATTTCTGCA GCTTTGCCTC TCCTGAGCGG AAGCCCCGTG AAGGCAAGGG
 151   AGCTAGCTGA TGGACTCTTT GTGGTCTTCT TCCTCTTTGC TCTGGAGACC
 201   CAACCAGGTG TTCTTAGGGG AAGGAGCACG TGAGTAGCCA AGAGGCTAAA
 251   AGCTGGTTCT CCCACATTCC AGGGTAAGTG ACTGGGTAGA GGGTGTGTCT
 301   GCCTCAGGCT GCTTGGAGGA GGTCCCCTGA AGGGCCATGA GAAAATCCTA
 351   CCCAGAGCCC TTGGTTTTCC AGCAGCCCTC CACCTAGAGG AAAGGAGCCT
 401   GTCGTTCTGA AGATGAAGAG TGGAGCCTAT GGGGGTGGGC AGATTGTGTC
 451   CTGGGACAAT GGGGTACCTA GAAGAGAAAG GAATCTCCTT TCGTTTGAGG
 501   TCTACCTGGG GGTCGTGTGT CTGTAAATGG GGTGGAGAGA GGAGAAGACA
 551   CAGATCTTAT AACGTAGATG CAGGAAATGC TGACAGTTCA GTGTAGAGAA
 601   CTTACTCAAT TCATATAGCC TCCAAAGCTA TCTCCTCAGG CAACGCAAAA
 651   CAAACCAGTT GGAGCCGCAA GACATCTAAT GGCTTATCGA GTCCCACACC
 701   CTCGATTCTT TGCTAATTTT ATGGTTTTGC TTTTGAGACA ATCTACTGTA
 751   GCCTAAGATA GCCCCAAACT CAAATGTAGC TGAGGCTGAC TGACCCTGAG
 801   CTCTGGAATT CCAGACACAT GCATATCTTT GCTAGGCAA TAATCGCTCT
 851   ACCAGCTGTA CTCCCACATT CCAGGGTAAG TGACTGGAAT TCTCACTTAC
 901   TATATCCCTT TAAAAATTCC CTGAGTGGGA TGGTTGTAGC CAGAGGGAAA
 951   AGGCACCAAC AACTGCTTGT CACTTTCCAA ATTTGGTAGC CTGAACAAAC
1001   CACTTATCAA GACAACAACT ATATATCATT TCTTTTCTTC TCTCTCTCTC
1051   TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTTTN
1101   GAAAGAGTCT CACTACTATG TAGCCCTTGA TAACCTAGAA CTCACTATGT
1151   AGTCCAGGCT TGGCCTTCAG CTCGCAGAGG TCCACTTGCC TTGGGAGTTG
1201   AGAGATTAAA GGGATGCATC TCCACATGTG TCCAACAGTG CTTTTTAAAA
```

3

```
1251  ATATTTTTAA AACCATGCTT ACAGCCAGGC ATAGTGGGCG TGCCTTTAAT

1301  CCCAGTACTG GGGAGGCAGA GGTAGGTAGA GTTCTGAGTT GGAGGCTAGC

1351  CACATAGTAA GTCCCAGGAT AGCTAGAACT ATGTAAAGAC CATGTCTCAA

1401  AAAAGATGCA CACACACATA TACACACACA CGTTTGTATG TGTTTGTTTA

1451  GTGTGTATGT GTGTGTACAC TTGCACATAA AGGTCAGAGT ACCACATTAC

1501  AGGAGTCAGT TTTCTCCTTT TATCATGTAT GGATGGAACA CGGGTCCATC

1551  CATAGCATCC TTAGCAGCAG GTATCCTTAT CCACTGAGCT ATCTCAGCAG

1601  CCCCACATTG CTTATTGGAT GTTTTTGGAT GAGGATAGTT ATATTAAAAA

1651  GGTTTCTGGT GTTGGTCTGG GTAGTTACCC TTTAACCCAT CTCTAGAGCC

1701  TGTCTCTTGA GTTTGAGGCC AGCCTGGTAT ATGTAGCTAG ACAAAGTTTC

1751  AAAAATGAAC AGAATCCTGG GACTAGAACC CATTTGTAGA ATGCTTGCAT

1801  AAGAAGCTCT GGGTTCAACT TCCTGCATCT CCAGAGGGAT TTTGTTCTGT

1851  AGTTTTAGTT TTTCAAGACA GAGTTTCTCT GTGTAGCCCT GGCTGTCCTG

1901  GAACTCACTC TGTAGACAAG CTGGCCTCG AACTCAGAAA TCCTTCTACC

1951  TCTACTTCAG GAGTGCTGGG ATTAAAGATG TGCGCTGCCC TCCTCCACCC

2001  CAATTTGTTT TTGTTTTTTA AGGGCCCCGG TAAACAGTAA ATTAACATGT

2051  GCATCCTGTT TGTCTTTGTA ATGACTCAAA TGTTGGGCTT CTGACCACTA

2101  GAGGGCAGCA GGCAGATACT AATGGACTGG GCGGAGAGAA GGGTAATCAG

2151  GAGCAGACCA GACTCGCGGA TAAACCAAAC AGCACCGCCA GCCGACCCTA

2201  GGCGAGGAGA GCGCCACAGG CACCAAGGGA AGACTTGAAG TAGTGTCTGA

2251  TCTCTACCGC TTCAGCAACC ATCGCGTTTG GGTGGGCTCC AGACAGGCAA

2301  AGTGCCAGCA AATGGTCCCT GTAGCTGACT AAACAGACTA TCAGACCCAA

2351  ACCACCACTG GACCGTGAAT GTTGCCCAGT GTGTTGCCTA GCCGCTTTCA

2401  GAATCCCAGC TTCTGGGTGT TGTGGAGGAA ACCCCTTAGC CTCGGTAACT
```

4

```
2451  TTCACCAGGC CCTTCTTGTC TCTAGACATC TAGACAGTTG GAAGCATCAG

2501  TCTTGACCCA GCCACCGGTT CAGATTCTTT GCCTTGCTTT TTCTTCCCCA

2551  GTTCAGCCCT GGCCAGGCCC CCAGGAAGAA TTTCCAGGGC CAGAGGGCAG

2601  CCTAAGGCAC AGATGCCCAC CCCTGCAATG TTCCCGCCAC ATGCCCAGTT

2651  CAGTACCCAG GGCCCAACCC CAGAGGGTGC GGAATGACAG ATTCTGACAA

2701  TCATTAAACC AGCCAGGCCT GATTCCCAG CACCGCCCGT TAGGATATGG

2751  GCCAAGTGGC ACGGAATATG CAAATCACAT GGGACAGGGA GCCCAGTCTG

2801  AAGGCCAGGA AATCCCCAGC ATCCAATGAG CCACCAGCTC AGGTTACAAC

2851  CGGGGACGTA CGCCGAAGAC CTGGAGGGGA GGAGCTCCTG CTTTGCTCTA

2901  TTTAGAGCGG GTGGGGGCAG CGCCCTGGCC ACACTCATCA CTGCTACCTG

2951  CGGAGCCTTC GACCGCTTAG ATTTTTTCCC TTCCTGTGGC CTCAGAAGCC

3001  TGCTCACCCG CCTGCCACAC CGAACCCTGA CACACCTCGG TACGGTCACA

3051  TTCCCTAACT CTGGCTCCAG GAACCGTCCA GTGGATTTAC AGTTCTGAAC

3101  TTAATCACTC AGGCTTGGAG GTTCCTAGCT GGAGTGTTGG GGCTACTGTG

3151  GGTGTATTCT GGGACTGGTC AGAGACCAGA TCGGTGTCTT GGAGGGACAG

3201  GGTGGCTTCT TTGGTTCAGG AGCCCACGTG ATTTGTGGAG AGACCCCAGA

3251  AGAATTTGTA TCATGCTCCC ACCCGCTTTG AGATTATTTT TATTTTTCGG

3301  AGCCGAATTT CCCAGTTTGG CGCCAGCTGG CCTGACCCTT CTAGGCTCAA

3351  GAGAGCATCC AACCTCAGCT TCCCCAAGTA GCTGGCTCTT GGTGGTGATG

3401  GTGGTGGTGG TGGTGGTGGT GGTGGTGGTG GTGTGTGTGT TTGTGTGTGT

3451  GTGGTGGTGG GGGGGGTGTT GAAGAGAGAA TGTCTACAGC AACACTGAAC

3501  TTCCTGCCTC TCGGCTGTTG CTGCCCAGGC TTTGCCAGAC AGAAATGGAA

3551  GTGTATCCTG ACCTGTACCC TCCCCACCTT GTCTCCTCTT CCCAGGGGCC

3601  CTCATGGCTG CCACCTGTGA GATCAGCAAC GTTTTTAGTA ACTACTTCAA

3651  CGCCATGTAC AGCTCAGAAG ACCCCACCCT GGCTCCTGCT CCTCCGACTA
```

5

```
3701   CCTTTGGCAC TGAAGACTTG GTGTTGACCC TGAACAACCA ACAGATGACA

3751   CTGGAAGGTC CAGGTGAGTG CTGTGTAAAA TCTTTTCAGA CAGGACACCA

3801   ATGATCTGAG AGGCTCTTAG ATGATAAATG GACAGGGAGG AAGGGTATCC

3851   TGGAGTTAGT GGCTGGGGAG GATTTATTCA TTCATATGTT TGTGTAGTAC

3901   TGGGGAAAGA ACCCAAACAA GACCTTATTT ATGCTAGACT GTGTTCCTAG

3951   TCCCGAGAAG ACTGTACTGG CTGAGGTGGT GGGAATATAA GAACTGTGGT

4001   GACAGATTAA GGGAGGATGA ACTTGAGAAC TAGCCATGTT GTGATTGTGG

4051   ATATGTATCT GTCCCTCTCC GCCCCTCCTC GGGTTGTGTA GGACCTCAGA

4101   CAAGATCCCA AAGGGACAGG ACTGATCCTC TGGCTGTACT CCACCTTGCA

4151   GAGAAGGCAA GCTGGACTAG CGAGCGGCCC CAGTTCTGGT CGAAGACCCA

4201   GGTTCTGGAG TGGATCAGCT ACCAAGTGGA GAAGAACAAG TATGACGCCA

4251   GCTCCATCGA CTTCTCCCGC TGCAACATGG ACGGAGCCAC CCTCTGCAGC

4301   TGTGCGCTGG AGGAGCTGCG GCTAGTCTTT GGACCTCTGG GAGACCAGCT

4351   CCATGCCCAG CTTCGGGACC TCAGTAAGTC TAGGCTGGGA GCCACAGGGC

4401   CTAAAGAGTG AGCGAGGTGG CTGGGACTTG GGCAGGAGGG TGCAGCCATC

4451   GAGCCCCTGC CGGAACCATG GTCGGTGACG CTCTCCCTCC CTGCCTCCGC

4501   CAGCCTCCAA CTCTTCTGAT GAACTCAGCT GGATCATCGA GCTGCTGGAG

4551   AAGGATGGCA TGTCCTTCCA AGAGAGCCTA GGCGACTTGG GCCCCTTTGG

4601   TGAGAACCCA TTTTCTCCCT TTTTCCTCCC TAGCTTGTCT TGTCCCATCT

4651   GTAACTCCTC CAGAGTGCTA CAGATATTCT CTCCCAACTT GAAAATAAGT

4701   CCATAGTCAT TTCTGTGGTC CCTGGAGGGT CGTGCCTGTC CTTGCTGGTA

4751   TCCTGGGCCT CTCTAAGCTC TTAACTTCTT TTCTCAGATC AGGGAAGTCC

4801   TTTTGCCCAG GAACTCCTGG ATGATGGCCG CCAGGCCAGT CCCTACTACT

4851   GCAGTACCTA TGGCCCTGGA GCGCCCTCCC CCGGCAGCTC TGATGTCTCC

4901   ACTGCAAGTA AGTCCTGCCC TTGCCACAGC CTGCCTTCTC CAAGTGCCCT

4951   AGAGTGCATC GAGTTCTTAC AATACTCATT CAGTATCTGA AGTCTGGGTA

5001   CGCAGTGACT GGGTAGGCTG GCCCTGGCAT TCAAGTGGTA TTCTTCACCC

5051   CTAGGGACCG CTACTCCCCA GAGTTCCCAT GCCTCTGACT CCGGTGGAAG

5101   TGATGTGGAC CTGGACCTCA CCGAgAGCAA GGTCTTCCCT AGAGGTGAGT

5151   TGAGGGCTGT TCTTGGGGGT CCTGTCCATG GGGTCTAGCC ACTCCCCTCT

5201   GCCCTATGGC TGCAGTTTCT GTACCAAGGC TCCCTGTTGA CACCCTGCCC

5251   TTACCTTCTC TTGACCTTCC AACCCCCTTC CCATAGATGA CTTTACTGAC
```

6

```
5301   TATAAGAAGG GGGAACCCAA GCACGGGAAG AGGAAACGGG GGCGTCCCAG

5351   AAAGCTGAGC AAGGAATACT GGGACTGTCT GGAGGGCAAG AAGAGCAAGC

5401   ACGGTAAGCT CTAAGGGCTG CCAGGCCTGT GGGCGGAGGG ATACTATTCC

5451   TTCAGCTTCC ACTGGCCTCT CACAGCCGCT GGAACTCATT GCATTGACGG

5501   GGCTCCATGG CATTTGTTAC TGCCTCTTTA CAGAGGCCTG CTTGGACTTA

5551   GAGAAGAAGG GAACTGAGGT CCTAGGAGAG GCCATGGAGA GAGTCCAGCC

5601   TTCCCACATT CTTCCTCTTT AACTATCCCT GTACTTGGCC CCTGTCTTGC

5651   CTGATGGAAC TTTCCGATGG GGAGGAGGCA GCTGGTGGGT GCTCAGAGCC

5701   AGGCAGGCTG GGGTGGCTGC GACTCCAGGC ACAGCCTGCT GAAAGAGCCT

5751   CGGGCCTTGT GTGCTCCAGC TGCTCTGAAC CACCTGGAGG TCACCCCAAA

5801   GGTCTTGCCT GCCGCCTCTT CCATACACAC AGACCTGCAC GTATGCACTC

5851   CTGGCCTGCT TTCCTTCTTG TCCCGAGGGG ACTGTTAATT CCGGGAAGCT

5901   GTTTCTTGGT CCCTCAGGCT ATAGGCAGCT CTCTGACCCC ATGTGTGCCA

5951   AGTTCTCACC ACCACTGGTC CCCACTGAAC CATGAGCCCC CTCACAAAGA

6001   AGCGTGTCTC TGTCGCTGTC CATCTTAACC AGTTGTTTGA TCCTTAACTG

6051   GTGAGAGAAT CGAGCGCTCT GTGCAGTCGG CCTAGCGCAT TGCATTTTGG

6101   GGCAGGAAAG GAAGCAGCCA CTATAGCAAT CACTAAGAGG ACATTTCATA

6151   TACTCCCATA TGCCTTGGCT CTTAGCCTCG TTGGGATAGG AGAGGCCAGG

6201   TCGCCTAGAG GAGAGGGGCA CCCCAGACTG ATAACTGAGG AAATCTTCCC

6251   TTGTAGCCCC CAGAGGTACT CACCTGTGGG AGTTTATCCG AGACATCCTA

6301   ATCCACCCCG AGCTCAACGA AGGCCTCATG AAGTGGGAGA ACCGGCACGA

6351   GGGTGTGTTC AAGTTTCTTC GCTCAGAGGC CGTGGCCCAA CTCTGGGGCC

6401   AGAAGAAGAA GAACAGCAAC ATGACCTATG AGAAGCTGAG CCGAGCCATG

6451   AGGTGAGTGT GAGCGTCAGG GACCTCTGCT TGGGCTCTAC TGGCTTCCGC

6501   TAGGTTTCAC GAGACAGGCC TGAGGCCCGT ATGGAGAGGA CAAGGACAGT

6551   GTTGTGGCCC TGTGTAGTTG GTTACGTGCA GCATGAAGAA AGCGCTGGGC

6601   AGAGATCGTG AGCACACTTA GCTTTAGCTA ACATTTCTGT GTTTCCTGCA

6651   GACTTGTTCT AAGAAAGACA CTTGAGAGAG AGAAAGAATA GAAATTGACA

6701   GCTCAGCTCC CTTGTCTCTG GGCCACAAAG GTGAACTAGC TCAGCATTGC

6751   TAAAGTCCCC TCTCCCTCAG TTCACGGGCC TTTATGAAAA GCCCCAGGAC

6801   ATAGCCAGAA GGCACAGAGA AGTAAATGTA GAAGCAGGTG CTCTGGCCAT

6851   AATTACAGAT CACCGCGGCC ACAACAGGTG AGGAGaGGGA ACACTCAGGC
```

7

```
6901    AGAGAGGGCC  AGCTCAGCAC  ACTGGGGCTG  GGAACCAATG  CGAACCTCAG

6951    TCCATAGCAT  GCCTCTTGCC  TACACCTCTG  ACCACCTCCT  TCCCACGCAG

7001    GTATTACTAC  AAACGGGAGA  TCCTGGAACG  GGTGGATGGC  CGACGGCTCG

7051    TCTACAAGTT  TGGCAAGAAC  TCTAGTGGCT  GGAAGGAAGA  AGAGGTTGGA

7101    GAGAGTCGGA  ATTAAGGATC  GGGGCTGGAC  CCAGGACCTG  ACTCAGGCAT

7151    GAACTCCAGA  ACTGAAGCCT  TCCTGGAAGG  ACAGGCAGGC  CTGACGGCCC

7201    CCTTAACATG  GATGTGTTCC  CTGTGTTGCT  GTAGAGAGGA  AGAACCTGTT

7251    GGGCGTGCCC  TCTGCAGTCT  CCTCAAGTGC  AGCCTTTGGC  CTCTCTCCTC

7301    GCCCTCTTGG  AATTACAAGC  CCCGGGTTTG  AACCAACTTG  TTCGATAACT

7351    CTTCCAGCTG  TGATTCCAGT  TCCCTCCCGT  CCCAACATGG  ACTGCAAATG

7401    AGACCCACCT  GCAGATGCCT  GGCCTCAGCC  AAGGAGGCTG  GGGAGACTGT

7451    GGCAGGAGAC  TGCAGGGACG  GAGGGGACAG  GGTTGTGTCC  TCGGTACTTC

7501    CTGGACTGCC  TTCCACCTCT  TTGCTCAGTA  CTCAGGCTCC  ACAGACGGGG

7551    GTCGGATCAT  CCCTAATTTA  TGTGCTATAA  ATATTCCAGG  TGTATATAGA

7601    GAGCTATTTT  TTCTAAAGCA  TTTCCCCTCC  CTGCTCTTCT  CCACTGAGTG

7651    CTGGTGGCCA  GACTGATTTT  TTTTTTAGCC  CCCCTAACTG  GACCAGCGAG

7701    AAGTAGGGTG  ATTCCAGGAC  CCCCTCTTCC  CCCAGAGGGG  TCTCCTGGAT

7751    CC
```

8